PROCEEDINGS OF TDWG

Biodiversity Information Science and Standards 1: e20382
doi: 10.3897/tdwgproceedings.1.20382

OPEN ACCESS

Conference Abstract

Three years of Xper^3 assessment: towards sharing semantic taxonomic content of identification keys

Amélie Pinel‡, Sylvain Bouquin§, Estelle Bourdon§, Adeline Kerner|, Régine Vignes-Lebbe§

‡ UPMC Univ. Paris 06, Sorbonne Universités, Paris, France
§ Institut de Systématique, Évolution, Biodiversité ISYEB - UMR 7205 – CNRS, MNHN, UPMC, EPHE UPMC Univ. Paris 06, Sorbonne Universités 57 rue Cuvier, CP48 F-75005, Paris, France
| CNRS UMR 7207, MNHN, Paris, France

Corresponding author: Régine Vignes-Lebbe (regine.vignes_lebbe@upmc.fr)

Received: 16 Aug 2017 | Published: 17 Aug 2017

Citation: Pinel A, Bouquin S, Bourdon E, Kerner A, Vignes-Lebbe R (2017) Three years of Xper^3 assessment: towards sharing semantic taxonomic content of identification keys. Proceedings of TDWG 1: e20382.
https://doi.org/10.3897/tdwgproceedings.1.20382

Abstract

Xper3 is a collaborative system that manages structured descriptive data on taxa or specimens. It is available online and linked to web services including two services for identification: a free (multiple) access key (Vignes Lebbe et al. 2015) and single access key (Burguiere et al. 2013). These web services use the TDWG Structured Descriptive Data format (SDD) (Hagedorn et al. 2005). The Xper3 platform was launched in November 2013. Three years later, 1990 users had created accounts and edited 2499 knowledge bases (KB). Unfortunately, there exists no public overview of the existing content. Each KB is autonomous and can be published as a free access key (e.g., http://cochenilles.bio-agri.org/mkey.html).

KB owners are free to publicize their keys in publications (Padovan and Magenta 2015 ; Engel et al. 2016) or on websites (http://acrinwafrica.mnhn.fr/SiteAcri/Xper.html)(http://herbaria.plants.ox.ac.uk/bol/caricaceae/Key).

This has two consequences:

- possible duplicate content or overlapping effort (e.g., several keys on orchids)

© Pinel A et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

- characters and states, documented by texts and images cannot be used for building another key without making copies.

In order to solve the first problem, we analyse Xper3 metadata (e.g., name of KB, owner, number of contributors, date of creation, date of last modification) and we provide an overview of the existing content. Firstly, 48% of KB are empty or inactive with extremely limited content (fewer than 3 taxa or characters), which has not been accessed for a long time. We discard these KB in our analysis and only consider the 1300 active KB. We also discard “test” KB and duplicate KB. Surprisingly, we discovered 15 medical KB for diagnosis of various diseases and 34 non-taxonomic KB (e.g., wine, fashion, computing equipment). For taxonomic KB, we present the taxonomic and geographic distribution of the KB (angiosperms and arthropods are prevailing taxa) and we compare with the number of known species.

The second point concerns semantic data sharing. We compute the rate of terms (character, character state) duplicated in several KB in the same taxonomic groups, in order to evaluate the interest in sharing resources between Xper3 KB. Then we look for existing ontologies in the bioportal (https://bioportal.bioontology.org/ontologies). Although Xper3 may manage structured data, its data model does not use an ontology language like RDF*1 or OWL*2 (ontology languages and tools offer unique identifier and reasoning mechanisms), and each KB has its own vocabulary.

We will be contacting KB owners to obtain more detailed metadata and to facilitate the automatic publishing of authorized KB on the Xper3 website. We plan also to implement an easy link between Xper3 and external ontologies to help editing new KB and to export KB data models to existing ontologies.

Keywords

Xper3, knowledge base, identification keys, ontology

Presenting author

Régine Vignes Lebbe

References

- Burguiere T, Causse F, Ung V, Vignes-Lebbe R (2013) IKey+: a new single-access key generation web service. Systematic Biology 62 (1): 157-161. https://doi.org/10.1093/sysbio/sys069
- Engel J, Brousseau L, Baraloto C (2016) GuiaTreeKey, a multi-access electronic key to identify tree genera in French Guiana. PhytoKeys 68: 27-44. https://doi.org/10.3897/phytokeys.68.8707

- Hagedorn G, Thiele K, Morris R, Heidorn PB (2005) Structured Descriptive Data (SDD) w3c-xml-schema, Version 1.0. Biodiversity Information Standards (TDWG). URL: http://www.tdwg.org/standards/116
- Padovan KD, Magenta MA (2015) Chave Interativa para árvores e arbustos da Área de Proteção Ambiental Serra do Guararú – Guarujá-SP. UNISANTA BioScience Vol. 4 nº 3: 160.
- Vignes Lebbe R, Chesselet P, Diep Thi MH (2015) Xper3: new tools for collaborating, training and transmitting knowledge on botanical phenotypes. Botanists of the 21st century.

Endnotes

*1 RDF: Resource Description Framework. This language of the semantic web uses triples "subject - predicate - value"

*2 OWL: Web Ontology Language